Description

METHOD FOR CALCULATION A PITCH PERIOD ESTIMATION OF SPEECH SIGNALS WITH VARIABLE STEP SIZE

BACKGROUND OF INVENTION

[0001] 1. Field of the Invention

[0002] The present invention relates to a method for calculating a pitch estimation, and more specifically, to a method for calculation a pitch period estimation of speech signals with variable step size.

[0003] 2. Description of the Prior Art

[0004] In the past few years electronic wireless communication has improved. At the same time the popularityof multimedia systems has increased while the demand for sound signal encoding and analyzing has become more and more popular. Sound telecommunication is an important application in the network of the next generation and has also an important role in multimedia telecommunications

in the network.

[0005] Telecommunication is widely applied to the techniques of sound signal encoding. So the telecommunication of specification is quite important. At the moment, there are some specifications of the International Telecommunication Union: PCM(64Kpbs)、G711(64Kpbs)、G726 (ADPCM、16、24、32、40Kpbs),G728(Low Delay CELP 16Kpbs)、G728(Low Delay CELP 8Kpbs). Currently, the cellular mobile telephone systems in North American use VSELP encoding techniques of the TIA(Telecommunication Industry Association). The cellular mobile telephone systems in Japan and Europe use RPE-LTP encoding techniques such as JDC(Japanese Digital Cellular) and GSM(Global System for Mobil Telecommunication). At the moment the current encoding technique is still at 8 Kbps. But the encoding technique of a new generation of mobile telecommunications is at 4.8 Kbps (LD-CELP)-2.4 Kbps (MELP,STC). For achieving such a ratio, the operation complexity is also raised, so that the general digital signal processor is used to finish the immediate operation.

[0006] For matching the design, there are digital signal processors in the special application design for sound compression or sound identification. The features of the DSP are: a

short instruction cycle, high parallelism and a plurality of special address modes to resolve the general digital signal processing.

[0007] The step with large amounts of operations in voice processing is the step of pitch estimation. This step is calculated according to equation 1.

$$R[\tau] = \sum_{n=0}^{N-1} x[n]x[n+\tau]$$

equation 1

$$pitch\ period = \{\tau \mid \max[R[\tau]]\}$$

[0008] Equation 1 is the operation of the autocorrelation. X[n]is a sound signal comprising a plurality of voice data from x[0] to x[N−1]. Voice data x[n+

τ :

] is a sound signal generated according to sound signal

x[n] which lags a lag parameter. The sound signal x[n+

τ :

] is from x[

τ

] to x[N−1+

τ

]. R[

τ:

is a autocorrelation value corresponding to a lag parameter. R[

τ :

] is the value that the amount of the voice data in the

sound signal x[n] times the corresponding voice data in

the sound signal x[n+

τ :

].

[0009] The autocorrelation operation in the method for estimating the pitch estimation, according to the prior art, calculates a plurality of autocorrelation value according to each lag parameter. Then a plurality of autocorrelation values are compared and the maximum autocorrelation value of these autocorrelation values are found. The lag parameter corresponding to the maximum autocorrelation value is used for calculating the pitch estimation.

[0010] Additionally, the normalizing autocorrelation method can also be used for estimating the pitch estimation. Please refer to equation 2.

$$R[\tau]^2 = \frac{[\sum_{n=0}^{N-1} x[n]x[n+\tau]]^2}{[\sum_{n=0}^{N-1} x[n+\tau]^2]}$$

$$pitch\ period = \{\tau \mid \max[Rn^2[n]]\}$$

[0011]   The normalizing autocorrelation method calculates the

value R[

τ

$]^2$ according to equation 2, i.e. the value R[

τ :

$]^2$ is calculated according to each lag parameter

τ

in a plurality of lag parameters

τ .

. The values R[

τ :

$]^2$ are stored in a memory and compared, until the maximum R[
mum R[

τ .

$]^2$ is found. Then a lag parameter

τ :

corresponding to the maximum R[

$\mathcal{T}$.

$]^2$ is used for estimating pitch estimation.

[0012] The amount of the operation of these two kinds of methods for estimating pitch estimation in digital signal processor is quite large. When the data bulk of the entry sound data is larger, the time of data processing is longer. When the sound signal cannot be operated immediately, the quality of the sound signal will be lowered.

SUMMARY OF INVENTION

[0013] It is therefore a primary objective of the claimed invention to provide a method for calculating a pitch period estimation of speech signals with a variable step size.

[0014] The claimed invention provides a method for calculating pitch estimation of a sound signal with a voice processor, the sound signal comprising a plurality of sound data, the method comprising the following steps:(a) providing an initial value to a lag parameter;(b) using the voice processor to calculate an autocorrelation value according to the lag parameter; (c) storing the lag parameter and the corresponding autocorrelation value in a memory; (d)setting a first increment and a second increment; (e)using the voice processor to compare the autocorrelation values in step (b) with a first threshold value, wherein when the autocorrelation value is less than the first threshold value, the lag parameter is increased by the first increment, and when the autocorrelation value is larger than the first threshold value, the lag parameter is increased by the second increment; (f)repeating the step (b), step (c), step (d) and step (e) until the lag parameter is larger than a predetermined value; and (g)comparing the plurality of autocorrelation values stored in the memory to find a maximum autocorrelation value and calculating a pitch estimation of the sound signal according to the lag parameter corresponding to the maximum autocorrelation value.

BRIEF DESCRIPTION OF DRAWINGS

[0015] Fig.1 is a block diagram of a voice processor according to the invention.

[0016] Fig.2 is a flowchart of a method for estimating a pitch estimation according to the invention.

[0017] Fig.3 is a flowchart of a method for estimating a pitch estimation in the first embodiment in the invention.

DETAILED DESCRIPTION

[0018] Please refer to Fig.1. Fig.1 is a block diagram of a voice processor 12 according to the present invention. A sound signal is an input in a voice processing device 10. The voice processing device 10 comprises a voice processor 12 for processing the sound signal $x[n]$, a memory 14 for storing a plurality of lag parameters and autocorrelation values R[

τ :

] calculated by the voice processing device 10 and a database for storing the sound signal x[n] and corresponding pitch range. The sound signal x[n] is generated by a sound signal generator 16 and input in the voice processing device 10.

[0019] Please refer to Fig.2. Fig.2 is a flowchart of a method for estimating a pitch estimation according to equation 1 in the invention. The method comprises the following steps:

[0020] Step 200: Providing an initial value to a lag parameter with the voice processor 12;

[0021] Step 202: using the voice processor 12 to calculate an autocorrelation value according to the lag parameter

τ :

; the autocorrelation operation can be operated according

to the above-mentioned equation 1 or equation 2;

[0022] Step 204: Storing the lag parameter

τ:

and the corresponding autocorrelation value R[

τ :

] in a memory 14;

[0023] Step 206: Setting a first incrementΔ_1 and a second incrementΔ_2;

[0024] Step 208: using the voice processor 12 to compare the autocorrelation values R[

τ :

] in step (b) with a first threshold value R_{th1}, wherein when the autocorrelation value R[

τ .

] is less than the first threshold value R_{th1}, the lag param-
eter

τ :

is increased by the first incrementΔ_1, and when the auto-correlation value is larger than the first threshold value R$_{th1}$, the lag parameter
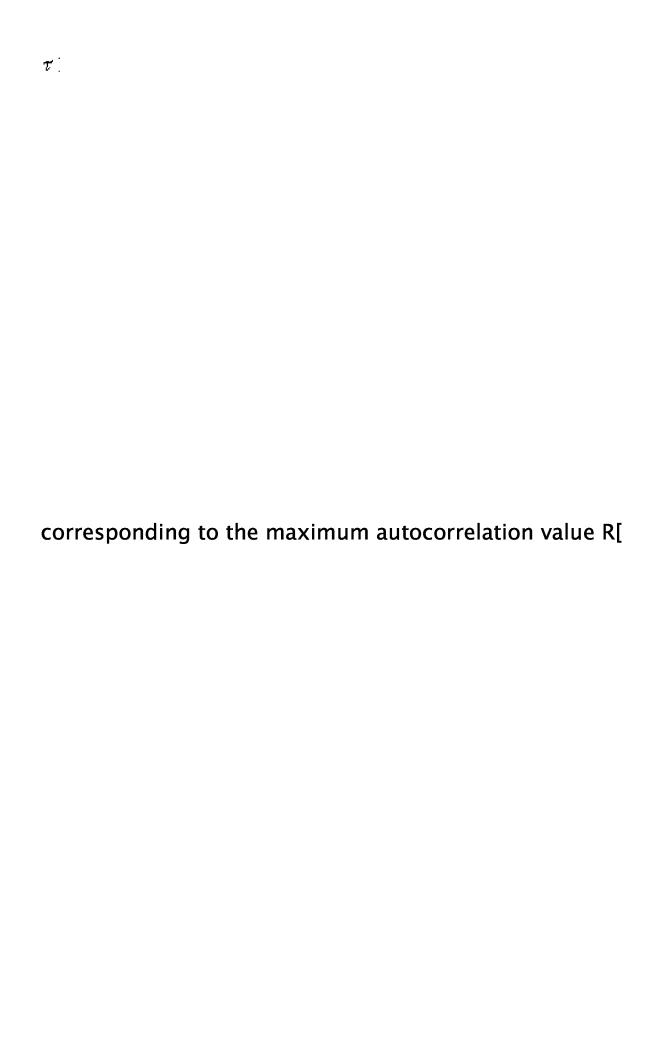
τ:

is increased by the second incrementΔ_2;

[0025]   Step 210: repeating step (b), step(c), step (d) and step (e) until the lag parameter

τ.

is larger than a predetermined value; and

[0026] Step 212: comparing the plurality of autocorrelation values R

τ :

] stored in the memory 14 to find a maximum autocorrelation value R[

τ :

] and calculating a pitch estimation of the sound signal according to the lag parameter

τ:

corresponding to the maximum autocorrelation value R[

τ .

].

[0027] In step 200 to step 204, the voice processor 12 is used for providing an initial value to a lag parameter

τ .

and calculating an autocorrelation value according to the lag parameter

τ .

. The lag parameter

τ:

and the corresponding autocorrelation values R[

τ :

] are stored in a memory 14. The initial value can be set as 1 or other value. In step 206 and step 208, a first increment Δ_1 and a second increment Δ_2 are set at first. The voice processor 12 compares the autocorrelation values R[

τ :

] in step (b) with a first threshold value R$_{th1}$. When the au-
tocorrelation value R[

τ .

] is less than the first threshold value R_{th1}, the lag param–
eter

τ .

is increased by the first incrementΔ_1. When the autocorrelation value R[

τ :

] is larger than the first threshold value R_{th1}, the lag parameter
rameter

τ :

is increased by the second incrementΔ_2.The incrementΔ_2 is less than the incrementΔ_1.

[0028] When the autocorrelation value R[

τ :

] is larger than the first threshold value R_{th1}, the lag parameter

τ :

is increased by the second incrementΔ_2. The purpose is to avoid ignoring the lag parameter

τ:

corresponding to the pitch estimation. When the autocorrelation value is larger than a first threshold value R_{th1}, the lag parameter corresponding to the autocorrelation value is close to the lag parameter corresponding to the pitch estimation of the sound signal and the second increment Δ_2 is increased by the lag parameter

τ.

. The second incrementΔ_2 can be set as 1 or other value that is less than the first incrementΔ_1. When the autocorrelation value R[

τ.

] is less than the first threshold value R_{th1}, the lag param-
eter

τ :

is increased by the first incrementΔ_1. The purpose is to ignore some lag parameters

τ .

to reduce the amount of the autocorrelation operations .

When the autocorrelation value is less than a first threshold value R_{th1}, the lag parameter corresponding to the autocorrelation value is not close to a lag parameter corresponding to the pitch estimation of the sound signal and the second increment Δ_1 is increased by a lag parameter

τ .

. The second incrementΔ_2 can be set as a larger value to ignore some lag parameters

τ

to reduce the amount of the autocorrelation operations. The first increment can be adjusted according to a different system.

[0029] In step 210, steps 202-208 are repeated. A plurality of autocorrelation values are calculated and stored in the memory 14 with a plurality of lag parameters. Because the autocorrelation is used for finding the level that the sound signal is similar to itself. When the sound signal is a cycle sound signal, the steps 202-208 are repeated until the lag parameter

τ .

is larger than the cycle number of the sound signal x[n].

When the sound signal is not a cycle sound signal, steps

202–208 are repeated until the lag parameter

τ

is larger than the number of the sound signal x[n].

[0030] The autocorrelation operation for the non-cycle sound

signal (ex: the noise or the sign) the autocorrelation val-

ues R[

τ :

] or the square of the autocorrelation values R[

τ :

$]^2$ cannot be used as the reference data for pitch estimation. Because the autocorrelation operation is used for finding the similar level between the sound signal and itself, a plurality of autocorrelation values of the cycle sound signal are showed in a regular pattern for finding the pitch estimation so that the pitch estimation can be found among the plurality of autocorrelation values. The autocorrelation values of the non-cycle sound signal are not showed in a regular pattern for finding the pitch estimation so that the pitch estimation of the sound signal cannot be found among the plurality of the autocorrelation values. In the embodiment, the autocorrelation oper-
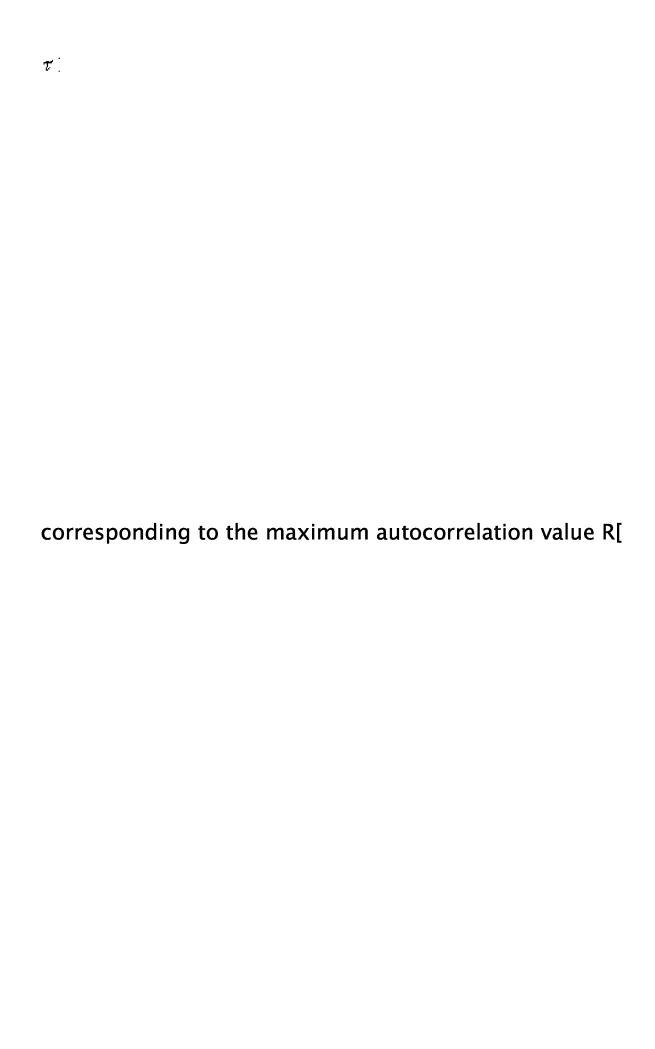
ation is only operated in the cycle sound signal to find the pitch estimation.

[0031] In step 212, the voice processor 12 is used for comparing the plurality of autocorrelation values R[

τ :

] stored in the memory 14 to find a maximum autocorrelation value R[

τ .

] and calculating a pitch estimation of the sound signal
according to the lag parameter

τ:

corresponding to the maximum autocorrelation value R[

τ.

].

[0032] The amount of the autocorrelation operations in the invention is less than the amount of the autocorrelation operations according to the prior art. The autocorrelation values are calculated according to each lag parameter

τ :

of a plurality of lag parameters

τ :

. The lag parameter

τ :

is increased by the first increment Δ_1 or the second increment Δ_2 in the invention. When the lag parameter

τ :

is increased by the first increment Δ_1 or the second incre-

ment Δ_2, the lag parameter between the lag parameter

τ :

and the lag parameter

τ :

$+\Delta_1$ or the lag parameter

τ :

$+\Delta_2$ are omitted. The autocorrelation values correspond-
ing to the omitted lag parameters can be set as zero or as
a smaller number.

[0033] In the invention, a third increment or a plurality of incre-
ments can be set. The autocorrelation values in the step
202 are compared with a second threshold value R_{th2}.

[0034] The second threshold value R_{th2} is larger than the first
threshold value R_{th1}. When the autocorrelation value R[

τ :

] is less than the second threshold value R_{th2} and larger than the first threshold value R_{th1}, the lag parameter

τ:

is increased by the second incrementΔ_2. When the auto-correlation value R[

τ :

] is larger than the second threshold value R$_{th2}$, the lag parameter

τ :

is increased by the third incrementΔ_3.

[0035]   Please refer to Fig.3. Fig.3 is a flowchart of a method for estimating a pitch estimation in the first embodiment of the invention. The embodiment is implemented in the voice processor 10.

[0036]   Step 300: Providing an initial value to a lag parameter with the voice processor 12;

[0037]   Step 302: using the voice processor 12 to calculate an autocorrelation value according to the lag parameter

τ :

; the autocorrelation operation can be operated according

to the above-mentioned equation 1 or equation 2;

[0038] Step 304: Storing the lag parameter

τ :

and the corresponding autocorrelation value R[

τ :

] in a memory 14;

[0039] Step 306: Setting a first incrementΔ_1 and a second incrementΔ_2;

[0040] Step 308: using the voice processor 12 to compare the autocorrelation values R[

τ :

] in step 302 with a first threshold value R_{th1}, wherein when the autocorrelation value R[

τ :

] is less than the first threshold value R_{th1}, the lag param-
eter

τ :

is increased by the first increment Δ_1, and when the auto-correlation value is larger than the first threshold value R $_{\text{th1}}$, the lag parameter

τ:

is increased by the second increment Δ_2;

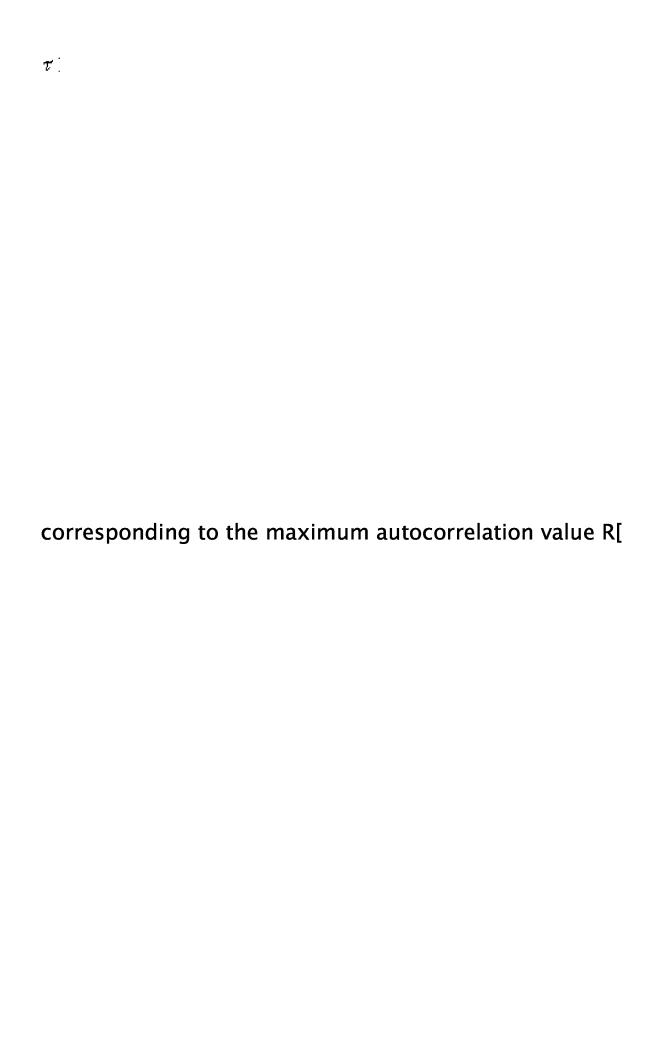[0041] Step 310: when the lag parameter

τ .

is larger than a predetermined value, step 312 is imple-

mented; when the lag parameter

τ :

is less than a predetermined value, step 302 is implemented; and

[0042] Step 312: comparing the plurality of autocorrelation values R[

τ:

] stored in the memory 14 to find a maximum autocorre-
lation value R[

τ:

] and calculating a pitch estimation of the sound signal according to the lag parameter

τ :

corresponding to the maximum autocorrelation value R[

τ .

].

[0043] The amount of the autocorrelation operations in the invention is less than the amount of the autocorrelation operations according to the prior art. The autocorrelation values are calculated according to each lag parameter

τ .

of a plurality of lag parameters

τ .

. The lag parameter

τ :

is increased by the first increment Δ_1 or the second increment Δ_2 in the invention. When the lag parameter

τ :

is increased by the first increment Δ_1 or the second incre-
ment Δ_2, the lag parameter between the lag parameter

τ :

and the lag parameter

τ :

$+\Delta_1$ or the lag parameter

τ :

$+\Delta_2$ are omitted so that the amount of operations can be reduced. And the lag parameter increases less for the second increment Δ_2 to avoid omitting the interval that the pitch estimation is probably in.

[0044] Those skilled in the art will readily observe that numerous modifications and alterations of the device may be made while retaining the teachings of the invention. Accordingly, the above disclosure should be constructed as limited only by the metes and bounds of the appended claims.